

University of Rhode Island

DigitalCommons@URI

Open Access Master's Theses

2013

Testing Theory-Based Quantitative Predictions with New Behaviors

Leslie Ann Daline Brick

University of Rhode Island, leslie.brick@gmail.com

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Recommended Citation

Brick, Leslie Ann Daline, "Testing Theory-Based Quantitative Predictions with New Behaviors" (2013).
Open Access Master's Theses. Paper 38.
<https://digitalcommons.uri.edu/theses/38>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

TESTING THEORY-BASED QUANTITATIVE PREDICTIONS WITH NEW
BEHAVIORS

BY
LESLIE ANN DALINE BRICK

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
IN
PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

2013

MASTER OF PSYCHOLOGY
OF
LESLIE ANN D. BRICK

APPROVED:

Thesis Committee:

Major Professor Wayne Velicer

Mark Robbins

Bryan Blissmer

Elizabeth Fallon

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2013

ABSTRACT

Theory development is essential for the generation and support of research ideas. Traditional Null Hypothesis Significance Testing (NHST) has been the *modus operandi* for testing research questions across many branches of science since the early 20th century. The focus of a statistical test under the NHST framework considers the rejection or acceptance of a null hypothesis based on a conditional probability of the data given that the null hypothesis is true (i.e. a p-value). This approach provides no direct support for a specific theory, which often takes the form of an alternative hypothesis. Furthermore, rejection of a null hypothesis based on a p-value provides no information on the magnitude of a difference and is affected by sample size, alpha level, and effect size. Such dependency on p-values can lead to misunderstanding and misinterpretation of results and conclusions. Therefore, the limitations of NHST warrant the investigation and development of new, more rigorous approaches to theory testing.

A quantitative approach, called “Testing Theory-Based Quantitative Predictions” (TTQP), has been proposed using effect size indices and confidence intervals to directly test predictions posited by theory (Velicer et al, 2008). Effect size indices provide information regarding the magnitude and direction of an effect while confidence intervals provide a means of “testing” specific predictions. This approach is an iterative process, allowing the researcher to tailor the theory as empirical data is collected. The use of the TTQP approach contributes to the movement away from NHST and the reliance on p-values, while promoting a stronger and more informative method. A quantitative orientation represents an essential change in thinking about theory testing by emphasizing the numeric strength of a measure, leading researchers away from a simple binary accept/reject framework. Predictions are made relative to the specific measure/variable, but the use of effect sizes allows for comparison across studies and

across theories. Therefore, the TTQP approach actually provides more information than traditional NHST.

The TTQP approach involves several steps. First, verbal descriptions of the expected values are designated a priori. These predictions are theory-based and guided by previous empirical findings (e.g. “small effect”). Second, verbal predictions are translated into quantitative values based on traditional guidelines or empirical results (e.g. “0.01”). Then, observed effect size estimates with surrounding confidence intervals are generated from sample data. If a confidence interval contains the predicted value, the prediction is confirmed. If the predicted value falls outside of the confidence interval, the prediction is not confirmed and explanations for failed predictions are examined.

The current study replicated findings from Velicer et al. (2008) and extended previous research by generating predictions for new health behaviors: diet and sun exposure. Secondary analyses were performed on cross-sectional data from a multiple health risk behavioral intervention. Predictions for each behavior varied slightly depending on the nature of the behavior and represented the major constructs of the Transtheoretical Model: decisional balance, self-efficacy, and processes of behavior change. Effect size indices were represented as ω^2 and 99% confidence intervals were generated to employ a stringent test of fit.

ACKNOWLEDGMENTS

The author would like to thank the members of her committee, Drs. Wayne Velicer, Bryan Blissmer, and Mark Robbins for their support and feedback, as well as the chair of the defense, Dr. Elizabeth Fallon. In addition, Drs. Colleen Redding and Joe Rossi offered crucial input and assistance regarding the Transtheoretical Model and theory testing. Finally, the students, faculty, and staff of the Cancer Prevention Research Center at the University of Rhode Island provided indispensable input on the topic during research meetings and expert panel discussions.

PREFACE

Manuscript format is used for the content of this thesis and is prepared for submission to the Psychological Bulletin.

The abstract was accepted to and presented at the Tenth Annual Society of Multivariate Experimental Psychology Graduate Student Pre-conference with co-author Wayne Velicer. A condensed version was published in Volume 47, Issue 6 of Multivariate Behavioral Research (2012).

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements.....	iv
Preface.....	v
List of Tables.....	vii
List of Figures.....	viii
Introduction.....	1
Method.....	13
Results.....	19
Discussion.....	21
Appendix.....	43
Bibliography.....	51

LIST OF TABLES

TABLE	PAGE
Table 1. Smoking Effect Size Predictions between PC, C, and PR at Baseline	28
Table 2. Diet Effect Size Predictions between PC, C, and PR at Baseline.....	29
Table 3. Sun Exposure Effect Size Predictions between PC, C, and PR at Baseline .	30
Table 4. Baseline Stage Distribution by Behavior	31
Table 5. Means (T-scores) and Standard Deviations of Smoking Variables by Stage.	32
Table 6. Means (T-scores) and Standard Deviations of Diet Variables by Stage.....	33
Table 7. Means (T-scores) and Standard Deviations of Sun Exposure Variables by Stage.....	34
Table 8. Quantitative Predictions and Observations of ω^2 for Smoking with 99% Confidence Intervals.	35
Table 9. Quantitative Predictions and Observations of ω^2 for Diet with 99% Confidence Intervals.	36
Table 10. Quantitative Predictions and Observations of ω^2 for Sun Exposure with 99% Confidence Intervals.	37
Table 11. Observed ω^2 Across Four Different Studies for Smoking Behavior.	38
Table A1. Quantitative Predictions and Observations of ω^2 for Smoking Behavior of Non-white Respondents with 99% Confidence Intervals.	45
Table A2. Quantitative Predictions and Observations of ω^2 for Smoking Behavior of Female Respondents with 99% Confidence Intervals.	46
Table A3. Quantitative Predictions and Observations of ω^2 for Smoking Behavior of Male Respondents with 99% Confidence Intervals.	47

LIST OF FIGURES

FIGURE	PAGE
Figure 1. Theory testing using quantitative predictions of effect size: comparison of predicted and observed ω^2 effect size value for Transtheoretical Model variables surrounded by 99% confidence intervals. From Velicer et al. (2008)	39
Figure 2. Comparison of predicted and observed ω^2 effect size values for smoking variables surrounded by 99% confidence intervals.....	40
Figure 3. Comparison of predicted and observed ω^2 effect size values for diet variables surrounded by 99% confidence intervals using behavioral stage criteria.....	41
Figure 4. Comparison of predicted and observed ω^2 effect size values for sun exposure variables surrounded by 99% confidence intervals using behavioral stage criteria.	42
Figure A1. Comparison of predicted and observed ω^2 effect size values for smoking variables of non-white participants surrounded by 99% confidence intervals.	48
Figure A2. Comparison of predicted and observed ω^2 effect size values for smoking variables of female participants surrounded by 99% confidence intervals.....	49
Figure A3. Comparison of predicted and observed ω^2 effect size values for smoking variables of male participants surrounded by 99% confidence intervals.....	50

“Testing Theory-Based Quantitative Predictions with New Behaviors”

by

Leslie Ann Brick¹, Wayne F. Velicer², James O. Prochaska²

is prepared for submission to Psychological Bulletin.

¹ Graduate Student, Behavioral Science, Cancer Prevention Research Center,
Department of Psychology, University of Rhode Island, Kingston, RI

² Professor, Cancer Prevention Research Center, Department of Psychology,
University of Rhode Island, Kingston, RI

Testing Theory-Based Quantitative Predictions with New Behaviors

Theory development is essential for the generation and support of research ideas. Traditional Null Hypothesis Significance Testing (NHST) has been the modus operandi for testing research questions across many branches of science since the early 1900's (Kline, 2004). Modern day NHST is the result of a blending of two schools of thought: the Fisherian approach, which simply features a statistical test of the null hypothesis; and the Neyman-Pearson approach, which introduces the alternative hypothesis, a fixed alpha level, specification of one- or two-tailed regions, as well as Type I and II errors (Kline, 2004). The focus of a statistical test under the NHST framework considers the rejection or acceptance of a null hypothesis based on a conditional probability of the data given that the null hypothesis is true (i.e. a p-value). This framework provides no direct support for a specific theory, which often takes the form of an alternative hypothesis.

Consequently, NHST tends to provide a result that is not what researchers are actually interested in: support for their theory or the alternative hypothesis (Kline, 2004; Steiger, 2004). Thus, the limited focus on a test of the null hypothesis has been criticized for almost as long as significance testing has existed. Furthermore, rejection of a null hypothesis based on a p-value provides no information on the magnitude of a difference and is greatly affected by the size of the sample, the alpha level, and the size of the effect. Such dependency on the p-value, which is arbitrary, can lead to misunderstanding and misinterpretation of results and conclusions (Cumming, 2012;

Kline, 2004). Therefore, the limitations of NHST warrant the investigation and development of new, more rigorous approaches to theory testing.

A quantitative approach, called “Testing Theory-Based Quantitative Predictions” (TTQP), has been proposed and uses effect size indices and confidence intervals to directly test predictions posited by a theory (Velicer et al, 2008). Effect size indices provide the researcher with essential information regarding the magnitude and direction of an effect while confidence intervals provide a means of “testing” the specific predictions generated by theory. By examining and comparing effect sizes, researchers can gain insight into the importance of individual measures or constructs within a theoretical framework. This approach is an iterative process, allowing the researcher to tailor the theory as empirical data is collected. The TTQP approach has been conducted on theory of smoking behavior, but needs to be replicated with new data and applied to predictions involving novel behaviors. The present study is a secondary data analysis that represents a replication of the Velicer et al. (2008) findings and an extension of the smoking based predictions to two new behaviors, diet and sun exposure. The use of the TTQP approach contributes to the movement away from NHST and the reliance on p-values, while promoting a stronger and more informative method of theory testing.

As mentioned above, TTQP uses effect size indices and confidence intervals to directly test predictions posited by a theory. Scientists conducting research to test a specific theory are often determining the adequacy of the theory given the data, thus an “accept-support” hypothesis testing method may be an appropriate and informative approach (Steiger, 2004). The goal of the “accept-support” method places emphasis on

supporting a theory and can be considered to have a focus on model fit rather than the rejection of a null hypothesis (i.e. the “reject-support” method).

Cumming (2012) describes TTQP as a model fitting approach: it allows for a comparison of how well a theoretical model can fit different sets of data. Due to the quantitative nature of the comparisons, incongruities between the model and the data are easily identified and subsequently examined. Ease of interpretation is furthered through the use of graphical displays of comparisons, such as Figure 1, displayed by Velicer et al. (2008). The graph contains each prediction for a list of variables in relation to a 99% confidence interval generated by point estimates obtained from empirical data. The variables are sorted by ascending effect size for easy identification of the variables with the largest effects.

The TTQP approach serves as a guide to generate specific quantitative predictions about the magnitude of the effect size for a certain measure or variable, while confidence intervals surrounding the observed effect size are employed to test the predictions. A quantitative orientation represents an essential change in thinking about theory testing by leading researchers away from a simple binary accept/reject framework. Previously, binary outcomes have been considered sufficient and the magnitude of an effect was largely ignored. This quantitative approach shifts traditional ways of thinking by emphasizing the numeric strength of the measure(s). Predictions are made relative to the specific measure, but use the use of effect sizes allows for comparison across studies and across theories. If a prediction lies outside of the observed confidence interval, it is not confirmed. Therefore, the use of effect sizes with confidence intervals actually provides more information than traditional

significance testing as the magnitude and direction of an effect will inform a researcher of the role specific measures play in the scientific theory.

The iterative nature of this approach allows the researcher to tailor the theory as empirical data is collected. Kline (2004) describes a “paucity” of replication in the literature of behavioral research. Application of the TTQP can build empirical cumulativeness by integrating and testing previous empirical results to new data. In addition, the TTQP approach promotes meta-analytic thinking by emphasizing effect sizes, confidence intervals, and the integration of previous empirical data.

Mathematically, two main components make up the TTQP approach: effect size estimation and confidence interval generation. Effect size predictions are conducted a priori and involve several steps. First, verbal descriptions of the expected values are designated for each prediction. These predictions, such as “small,” “medium,” “large,” or “no effect,” are theory-based and guided by previous empirical findings. Second, verbal predictions are translated into quantitative values based on Cohen’s (1988) guidelines. Then, if available, empirical results from a previous study are used to recalibrate quantitative values to better reflect the empirical findings. Finally, the effect size estimates and confidence intervals are generated from the new data set. If a confidence interval contains the predicted value, the prediction is confirmed. If the predicted value falls outside of the confidence interval, the prediction is not confirmed and explanations for failed predictions are examined. Visual graphics are utilized to aide the interpretation of results. Point estimates for multiple measures can be displayed together for easy comparison and examination of missed predictions.

Effect Size Estimation

Explicit effect size predictions allow researchers to gain information beyond the binary accept/reject decision procedure by moving the focus of research away from p-values. Traditional NHST does not emphasize the magnitude of a difference, thereby ignoring a crucial piece of information regarding the relationship between variables. Quantitative predictions of effect size estimates guide a theorist to numerically clarify what their theory predicts. These numeric predictions then allow for quantitative comparisons between groups. The proposed study will utilize omega-squared (ω^2), which is used when comparing differences across more than two groups. This metric also corrects for the positive bias of explained variance when estimating a population effect by taking random error into account. It is calculated using information from a one-way between-groups fixed effects ANOVA:

$$\omega^2 = (SS_{\text{BETWEEN}} - (k - 1) * MS_{\text{WITHIN}}) / (SS_{\text{TOTAL}} + MS_{\text{WITHIN}})$$

where SS_{BETWEEN} and SS_{TOTAL} are the Sum of Squares between groups and Sum of Squares total, k is the number of groups, and MS_{WITHIN} is the Mean Square within groups. By directly testing a theory with quantitative comparisons, such as ω^2 , researchers can come to more precise conclusions.

Cohen's (1988) guidelines for the interpretations of variance-accounted-for effect sizes suggest a "small" effect to be one percent of the variance-accounted-for, a "medium" effect to be six percent, a "large" effect to be fourteen percent or more, and "no effect" to be zero percent. By verbalizing effect size predictions, comparisons can be viewed with greater practical understanding, while assigning numeric values to verbal predictions allows the procedure to become quantitative. However, these

classifications are very broadly defined and were intended only as a guide to initial estimates. As theorists become familiar with the effects found in a specific area of research or population, they should adjust the classifications of “small,” “medium,” and “large” to better represent that area. Thus, for one study a medium effect size may be better represented by a classification of $\omega^2 = 0.08$ rather than $\omega^2 = 0.06$. Since part of this study examined two new behaviors, diet and sun exposure, Cohen’s original classification were used to guide effect size prediction estimations. As new data is generated for these behaviors, the effect size classifications may be recalibrated to reflect the data (Velicer et al., 2008). Depending on the population and context of the data, classifications may vary across studies.

Confidence Interval Generation

Confidence intervals surrounding point estimates are becoming more widely used as recommendations and formulas for calculation emerge. The *APA Publication Manual* recommends the use of confidence intervals whenever possible (APA, 2001). Cumming and Finch (2001) suggest four reasons for using confidence intervals: “They (a) are readily interpretable, (b) are linked to familiar statistical significance tests, (c) can encourage meta-analytic thinking, and (d) give information about precision” (p. 532). Use of confidence intervals gives us a richer, more dynamic sense of what our data can reveal.

Estimation of confidence intervals recognizes that sampling error may occur in point estimates generated by sample data. Intervals can be used to specify a range of plausible values at different levels of confidence. Use of a 95% level of confidence

will generate a narrower interval than a higher level, such as 99%, as it allows for slightly more error.

Calculation of confidence intervals around normal distributions is fairly straightforward; however, the calculation of confidence intervals around effect sizes such as ω^2 relies on the noncentral distribution. In fixed-effects between-subjects ANOVA, the noncentral F distribution is characterized by three parameters: the numerator and denominator degrees of freedom and the noncentrality parameter, lambda (λ). As λ increases, the distribution becomes more noncentral. In terms of hypothesis testing, the null hypothesis states that the noncentrality parameter equals zero. Thus, if a null hypothesis is false, λ will depart from the null distribution (Steiger, 2004).

Rationale for the movement away from traditional null hypothesis significance testing has existed for quite some time, but widespread use of confidence intervals has been hindered by the difficulty of computing exact intervals for noncentral distributions until the advancement of computer capabilities in the 1980's and 1990's (Smithson, 2001; Steiger, 2004). Kromrey and Bell (2010) provide a SAS macro that applies the interval inversion approach to the calculation of confidence intervals around ω^2 (Steiger, 2004; Steiger & Fouladi, 1997). When a distribution is not symmetric, such as the F-distribution when the null-hypothesis is false, the relative frequencies for λ change as the distribution changes shape, dispersion and location (Kromrey & Bell, 2010; Smithson, 2001). The interval inversion approach calculates a confidence interval around λ by using an iterative approach that seeks the values of λ that demarcate top and bottom percentiles of the noncentral distribution for a

particular F distribution. The upper and lower noncentral values can then be transformed into upper and lower ω^2 values. For additional information regarding the calculation of confidence intervals around effect sizes, see Fidler & Thompson (2001)

The Transtheoretical Model

The TTQP approach is used in the context of an explicit theory and provides an alternative framework to test predictions regarding the relationship between constructs described within that theory. In this paper, TTQP will be used to assess predictions of the relationship between health behavior constructs theorized by the Transtheoretical Model of behavior change (TTM; Prochaska & DiClemente, 1983). Briefly, the TTM is a theory of intentional behavior change and is used to assess an individual's readiness to engage in change. It utilizes a stage approach with three core constructs including decisional balance, self-efficacy, and processes of change. Behavior change is regarded as movement through a series of five stages: precontemplation, contemplation, preparation, action, and maintenance. Individuals across each stage of change utilize the TTM constructs in different ways, allowing for comparison of constructs by stage using effect size indices to gauge the direction and magnitude of differences.

The three behaviors examined in this study are smoking, diet, and sun exposure. The core TTM constructs for each of these three behaviors, decisional balance, self-efficacy, and processes of change, may be indicated by slightly different variables, depending on the nature of the behavior. For example, the self-efficacy constructs of smoking and diet are represented by a temptation scale, but is

represented as a confidence scale for sun exposure. Similarly, there are ten core processes of behavior change in the TTM, but a certain behavior may have additional processes specific to that behavior. For example, diet has one additional process, Interpersonal Systems Control (IS), and sun exposure has six additional, Health Care Provider (HC), Interpersonal Systems Control (IS), Reducing Exposure (RE), Regret (RE), Sunscreen Use (SS), and Health Responsibility (HT). Additionally, for sun exposure, Stimulus Control (SC) is left out, as this particular process does not have a theoretical basis for behavior relating to sun exposure.

Velicer, et al. (1999) conducted a longitudinal study that tested 40 differential a priori, TTM-based predictions. A large, representative sample of smokers in the state of Rhode Island were assessed and grouped by their current stage of change. Hypotheses were theory-driven with explicit quantitative predictions of the expected effect size of movement from one of three initial stages to stage membership 12 months later. Five variables representing constructs from the TTM (pros, cons, habit strength, positive evaluation strength, and negative evaluation strength) provided a series of comparisons involving stage transitions (progressing, stable, regressing). Thirty-six of the 40 predicted effect sizes were confirmed by the data, however, the study used significance tests instead of confidence intervals to support the hypotheses.

Velicer, et al. (in press) replicated the 40 predictions developed by Velicer et al. (1999) on a new sample of smokers using recalibrated effect size indices. This study combined the longitudinal approach developed by the 40 predictions with the TTQP approach using confidence intervals instead of significance tests. Thirty-two of the 40 predictions were confirmed and an analysis of prediction failure was conducted to

examine reasons behind a failed prediction. Most failures were attributed to the need for recalibration or sample fluctuation. Three of the failures are attributed to the Cons scale and require revision of theory. In these cases, the effect sizes predicted were larger than the effect sizes observed and were not contained in a 99% confidence interval.

A cross-sectional study employed the TTQP approach on a sample of smokers (Velicer et al., 2008). Fifteen effect size predictions were made that compared 15 variables within the three TTM scales (decisional balance, temptations, and processes) across the first three stages of change. Effect sizes were recalibrated using values observed in two previous smoking studies in order to better represent effect sizes in the area of smoking cessation research. Eleven of 15 predictions were confirmed and missed predictions were evaluated. Missed predictions were determined to be a result of four potential issues: sample fluctuation, need for theory revision, theory incorrect, or need for further calibration of effect sizes.

The current study replicated the findings of Velicer et al. (2008) and extended previous research by applying the smoking-based predictions to two new health behaviors. Secondary analyses were performed on cross-sectional data from a multiple health risk behavioral intervention (Prochaska et al, 2005). Study 1 directly replicated the smoking cessation results. Study 2 applied predictions to unhealthy diet and Study 3 applied predictions to sun exposure. Predictions for each behavior varied slightly depending on the nature of the behavior and represented the major constructs of the TTM: decisional balance, self-efficacy, and processes of change. Effect size indices

were represented as ω^2 and 99% confidence intervals were generated to employ a stringent test of fit.

Method

Study Design

Secondary analysis was conducted on baseline data from a population based multiple risk factors behavioral intervention. See Prochaska et al. (2005) and DePue et al. (2008) for an overview of the study. Participants were recruited and assessed for smoking, diet and sun exposure behaviors. Smoking behavior was assessed in regard to self-reported daily smoking habits. Diet was assessed in regard to a high fat diet: greater than 30% calories and total score on the Dietary Behavior Questionnaire (Greene et al., 1994; Greene et al., 1996). Finally, sun exposure was assessed in regard to self-reporting of exposure: 15 or more minutes of exposure per day or inconsistent SPF-15 use and total score on the Sun Protection Behavior Scale (Rossi et al., 1995).

Sample

A large health insurance organization provided patient information and 5407 primary care patients agreed to take part in the study. Participants were at risk for at least one of four behaviors: reduce smoking, improve diet, decrease sun exposure, and receive regular mammograms. Participant data from the baseline measurement was examined for three behaviors: smoking, diet, and sun exposure. The total baseline sample was 68.0% female, predominantly white (96.7%), 1.3% Hispanic, and had a mean age of 44.7 (SD = 12.7).

Measures

Demographics. Single items assessed age, sex, race, and ethnicity. Race and ethnicity were represented by two separate questions. One question asked for race (White, Black, Asian/Pacific Islander, American Indian/Alaska Native, Other) and a second question asked for Spanish/Hispanic origin (yes, no).

Stages of change. Participants were classified into one of five stages of change using an algorithm that assessed their readiness to change. For diet and sun exposure behaviors, stage is assessed based on an individual's perception of their readiness to change, and then adjusted based on a series of questions regarding their habits and behaviors to best reflect their readiness to change. Adjustment for these behaviors is due to the potential for discrepancy between self reported (i.e. perceived) readiness and an individual's actual behavior. The Dietary Behavior Questionnaire (Greene et al., 1996) is used to stage for diet and includes questions regarding an individual's intention to avoid eating high-fat foods followed by a behavioral assessment of that intention. For example, items include: "Do you sometimes eat fruit and vegetables as snacks" and "Do you eat reduced fat or low-fat cheese". Similarly, the Sun Protection Behavior Scale (Rossi et al., 1995) classifies participants regarding their intention to consistently protect themselves from sun exposure followed by a behavioral assessment of that intention. For example, items include: "Do you avoid the sun during the mid-day hours" and "Do you use a sunscreen with SPF of 15 or more on all sun exposed skin areas". For smoking, only one stage measure is assessed based on their readiness to quit. Participants were measured at baseline and grouped into one of the first three stages of change, representing the "pre-action" stages: precontemplation (PC), contemplation (C), or preparation (PR). A detailed discussion of stage of change

measures for smoking is discussed by DiClemente et al. (1991); for diet, see Green et al. (1994, 1999); for sun exposure, see Weinstock et al. (2002) and Rossi et al. (1997).

Decisional balance. Cognitive and motivational aspects of decision-making are measured by the Decisional Balance Inventory (Prochaska et al., 1994; Velicer et al., 1985). These two constructs are distinguished as the “Pros” and “Cons” of engaging in a behavior. For smoking and diet, the pros represent the perceived benefits of engaging in the unhealthy behavior and cons represent the disadvantages. For sun exposure, the reasoning is reversed. The pros represent the benefits of sun protection while the cons represent the difficulties or disadvantages.

Self-Efficacy. A person’s self-efficacy, or belief that they can prevent or cope with the temptation to fall back into unhealthy or high-risk behavior, is measured using measures of confidence and situational temptations (DiClemente, 1986; Velicer et al., 1990). These measures may be conceptualized and/or measured differently depending on the behavior of interest.

Situational temptations. The temptation to engage in negative health behavior in various situations was measured using the Situational Temptations Inventory (Velicer et al., 1990). This scale consists of three subscales that measure responses using a 5-point Likert scale (1 = not important to 5 = extremely important). For smoking, the subscales measure the Positive/Social, Negative/Affective, and Habit/Addictive aspects of the temptation to engage in smoking behavior. For diet, the subscales include: Positive/Social, Negative/Affective, and Difficult Situations.

Confidence. Confidence was measured using a scale designed to assess a person’s confidence to engage in sun protection behavior during difficult situations.

This scale has two subscales, general and sunscreen use, and is represented by a sum score of each subscale, with a higher score representing higher confidence.

Processes of change. The processes of change represent ten different behavioral and experiential strategies for changing behavior (Prochaska, Velicer, DiClemente & Fava, 1988). Experiential processes include: Consciousness Raising (CR), Dramatic Relief (DR), Environmental Reevaluation (ER), Self-reevaluation (SR), and Social Liberation (SO). Behavioral processes include Stimulus Control (SC), Counter Conditioning (CC), Reinforcement Management (RM), Self-Liberation (SL), and Helping Relationships (HR). These ten variables make up the “core” of the TTM processes of change, but additional processes specific to certain behaviors are often added to examine different concepts. For diet, one additional process is added, Interpersonal Systems Control (IS), resulting in a total of 11 processes. For sun exposure, six additional processes are added, Health Care Provider (HC), Interpersonal Systems Control (IS), Reducing Exposure (RE), Regret (RG), Sunscreen Use (SS), and Health Responsibility (HT). Stimulus Control (SC) is removed resulting in a total of 15 processes.

Statistical Analyses

Predictions were conducted by comparing participants classified by stage at a baseline assessment. Stage of change was considered the independent grouping variable while the decisional balance and self-efficacy subscales, as well as the processes of change, were considered the dependent variables. Comparisons between

stages were achieved by one-way between-groups fixed effects ANOVA. Information from the ANOVA source table provided the information necessary for calculating ω^2 .

99% confidence intervals for ω^2 were calculated using a SAS Macro (Kromrey & Bell, 2010). Predicted and observed values and upper and lower bounds for observed values were entered into an Excel spreadsheet that generates a graph allowing for a visual comparison of the prediction with the observed values.

Initial Effect Size Predictions

Effect size predictions for smoking have been examined previously (Velicer et al., 2008). Study 1 used the recalibrated effect sizes suggested in Velicer et al. (2008) as predictions in order to replicate and validate findings for smoking behavior. In the absence of previous empirical work for diet and sun exposure, initial predictions for Study 2 and Study 3 followed the original predictions of Velicer et al. (2008) that were guided by TTM theory and put forth before recalibration of observed data for smoking behavior. A few of the processes variables are unique to sun exposure and diet and therefore did not have initial predictions from the list previously tested smoking predictions. A panel of TTM, diet, and sun exposure experts met to discuss appropriate predictions for these behaviors and they were used for subsequent analysis.

Study 1: Smoking

Fifteen Effect size predictions for smoking were examined. Predictions represent a direct replication of results from Velicer (2008) and integrate the

recalibrated effect size predictions, therefore verbal predictions are not included. Table 1 displays the quantitative predictions for smoking effect size based on previous data.

Study 2: Diet

Sixteen predictions for diet (eating healthier) were examined. Verbal predictions based on TTM theory were translated into quantitative values based on Cohen's (1988) guidelines for interpretation of effect size. Table 2 displays the verbal and quantitative predictions for diet.

Study 3: Sun exposure

Twenty predictions for sun decreasing exposure were examined and are identical to those predicted for diet, except for the self-efficacy measure. For sun exposure, self-efficacy is represented as General Confidence and Sunscreen Confidence. Table 3 displays the verbal and quantitative predictions for sun exposure effect sizes.

Results

Baseline stage distributions for each behavior are presented in Table 4. Tables 5 displays sample sizes, means (represented as T-scores), and standard deviations by stage of change for Study 1. Table 6 presents samples sizes, means (represented as T-scores), and standard deviations by behaviorally corrected stage of change for Study 2 (Diet). Similarly, Table 7 displays samples sizes, means (represented as T-scores), and standard deviations by behaviorally corrected stage of change for Study 3 (Sun Exposure).

Study 1: Smoking

Thirteen of 15 predictions were confirmed for smoking behavior. Table 8 represents a summary of results for the 15 variables. The two misses include the Pros and SL variables. A value of $\omega^2 = 0$ was predicted for the Pros scale was predicted, but a value of $\omega^2 = 0.022$ [99% CI 0.004, 0.047] was observed. For SL, a value of $\omega^2 = 0.19$ was predicted, but a value of $\omega^2 = 0.101$ [0.046, 0.164] was observed. Figure 2 displays predicted effect sizes and observed confidence intervals for variables in increasing order of effect size. Misses are indicated by a dot falling outside of the upper or lower boundaries of the confidence interval.

Study 2: Diet

Six of 16 predictions were confirmed for diet behavior based on the behavioral criteria stage. Table 9 represents a summary of predictions and results for the 16 variables. The misses include: Pros, Neg./Aff., CC, DR, ER, HR, IS, RM, SC, and SL.

Figure 3 displays predicted effect sizes and observed confidence intervals for variables in increasing order of effect size.

Study 3: Sun Exposure

Five of 19 predictions were confirmed for sun exposure behavior based on the behavioral criteria stage. Table 10 represents a summary of results for the 19 variables. The misses included: Cons, Sunscreen Confidence, General Confidence, CC, ER, HR, HT, IS, RE, RM, RG, SR and SS. Figure 4 displays predicted effect sizes and observed confidence intervals for variables in increasing order of effect size.

Discussion

The iterative nature of the TTQP approach necessitates a close investigation of failed predictions for each study. By understanding why a prediction has failed to be confirmed, a researcher can modify and adjust expected values of effect size for future use, thus providing specific empirical values for constructs within a theory. As data accumulates across studies, the effect size values become true quantitative predictions rather than categorical values (i.e. small, medium, large).

Four explanations exist that may explain a failed prediction: sample fluctuation, theory revision, incorrect prediction, or further calibration needed. First, the use of 99% confidence interval permits a small number of misses due to chance fluctuations in a sample. These misses tend to be “near misses” and are very close to falling within the interval generated by sample data, but instead fall just outside of it. Thus, a future replicate may in fact confirm these near misses. Second, theory revision is required when an observation falls in the opposite direction of or very far away from the prediction. Such may be the case when a small effect was predicted and a very large effect was obtained. The theory made an inadequate prediction and need to be revised to predict large effects rather than small effects for that variable. Third, a prediction may be incorrect when observations are undoubtedly discordant with predictions such that the theory led to an overwhelmingly incorrect prediction. It may be the case that the theory itself needs major reconsideration to better reflect empirical observation, as opposed to a slight revision, as explanation two describes. Finally, further recalibration may be needed when observations and their confidence intervals do not align with any of the predicted values. This would be shown by variables that

demonstrate consistent effects that do not fall within Cohen's guidelines. For example, for a particular behavior, a medium effect size may be better represented by a value of $\omega^2 = 0.08$ rather than 0.06 if observed values fall more frequently near that estimate. It should be noted that recalibration is specific to the population and behavior of interest and should always be considered in regard to context; thus, a large effect for smoking may be presented as $\omega^2 = 0.19$ while a large effect for diet may be presented as $\omega^2 = 0.14$.

TTQP

Study 1 represents a successful replication of the smoking prediction from Velicer et al. (2008). Thirteen of 15 predictions were confirmed and the two misses were near misses, most likely due to sample fluctuation. It is evident that although the Pros and Self Liberation measures were not confirmed by the current study, they fell just outside the confidence interval but are still consistent with previous findings. Pros (observed $\omega^2 = 0.022$, [99% CI: .004, .047]) was posited to have no effect, but was observed to a lower limit of .004, which is essentially zero and thus can be considered a near miss. Self-Liberation measure (observed $\omega^2 = 0.101$, [99% CI: .046, .164]) represents a fairly large effect size but was found to be lower than predicted ($\omega^2 = 0.19$). Regardless, the upper interval is still quite large in the context of smoking variables and though sample fluctuation may be the reason for this miss, interpretation remains conceptually the same.

As data accumulates across replications, it becomes possible to generate true quantitative predictions. This study represents the third replication for smoking

behavior variables in the TTM. Table 11 displays a compilation of ω^2 values obtained across three studies. Data from the current study was compared to data from Velicer et al (2008) and one other study presented in Velicer et al. (2008), the Random Digit Dial (Fava et al., 1995). Though there exists some variation in the point estimates, effect size categories tend to remain consistent across studies.

Results from Study 2 indicate that predictions applied to smoking behavior do not directly transfer to predictions for diet. Six of 16 predictions were confirmed using the behavioral stage criteria. Five of the failed predictions may be considered near misses: Pros (observed $\omega^2 = .018$, [99% CI: .008, .030]), Environmental Reevaluation (observed $\omega^2 = .030$, [99% CI: .012, .054]), Helping Relationships (observed $\omega^2 = .030$, [99% CI: .012, .054]), Interpersonal Systems Control (observed $\omega^2 = .036$, [99% CI: .01, .061]), and Self Liberation (observed $\omega^2 = .101$, [99% CI: .067, .137]). Future replications may capture the predicted values given that observed values are very close to predicted values, which lie just outside the 99% confidence intervals. The remaining five failed predictions require theory revision. In general, predictions tended to be low, thus larger than expected effects were observed for many variables. Small effects were predicted for Negative/Affect (observed $\omega^2 = .057$, [99% CI: .039, .078]), Counter Conditioning (observed $\omega^2 = .070$, [99% CI: .042, .102]), and Reinforcement Management (observed $\omega^2 = .068$, [99% CI: .040, .100]), however, medium effects were found. Predictions should be increased to $\omega^2 = 0.06$ for each of these variables. Dramatic Relief (observed $\omega^2 = .112$, [99% CI: .077, .150]) was predicted to have a medium effect, but was instead observed to be large. The prediction for Dramatic Relief should be increased to $\omega^2 = 0.14$. Finally, one

prediction, Stimulus Control (observed $\omega^2 = .026$, [99% CI: .009, .048]), was too high and should be decreased from a medium effect to $\omega^2 = 0.01$.

Results from Study 3 indicate that effect sizes for sun exposure are strikingly different than those observed in smoking. Five of 19 predictions were confirmed using the behavioral staging criteria. An examination of the misses reveals that many of them are discordant with theory, signifying the need for theory revision and new effect size categories. First, three variables (Helping Relations, observed $\omega^2 = .035$, [99% CI: .016, .059]; Self Reevaluation, observed $\omega^2 = .101$, [99% CI: .069, .135]; and Health Care Provider, observed $\omega^2 = .038$, [99% CI: .018, .063]) can be interpreted as near misses due to sampling fluctuation. Helping Relations was predicted as $\omega^2 = 0.01$, and the lower bound of the 99% confidence interval just missed it at $\omega^2 = 0.016$. Similarly, Health Care Provider was predicted as $\omega^2 = 0.01$, and the lower bound of the confidence interval missed it at $\omega^2 = 0.018$. Self Reevaluation was predicted as $\omega^2 = 0.14$ and the upper bound of the 99% confidence interval just missed it at $\omega^2 = 0.135$. Future replication will reveal if these are sample fluctuations.

Next, theory revision can capture findings from another of the four missed prediction. Environmental Reevaluation (observed $\omega^2 = .126$, [99% CI: .091, .163]), Interpersonal Stimulus Control (observed $\omega^2 = .123$, [99% CI: .088, .160]), and Sunscreen Use (observed $\omega^2 = .135$, [99% CI: .099, .172]) were all predicted to have a medium effect. However, observed effect sizes were larger than predicted and should be revised as $\omega^2 = .14$ (large effect) for future studies. Finally, three misses can be revised as medium effects. Counter Conditioning (observed $\omega^2 = .050$, [99% CI: .027, .078]), Reinforcement Management (observed $\omega^2 = .056$, [99% CI: .031, .084]), and

Regret (observed $\omega^2 = .059$, [99% CI: .034, .088]) were each predicted to have small effects, but demonstrated medium effect sizes and would better be captured by $\omega^2 = .06$.

The remaining misses from Study 3 represent incorrect predictions and the creation of a new effect size category. The prediction for Cons (observed $\omega^2 = .170$, [99% CI: .143, .197]) fell more than one effect size class lower than the observed effect. It was predicted to have no effect, but was found to be a large effect and should be predicted as large in future replications. Many of the remaining measure observations also discordant with theory but revealed extremely large effect sizes, thus a new category of $\omega^2 = .23$ can be created. This “extra large” category would be adequate to capture four of the missed predictions: Sunscreen Confidence (observed $\omega^2 = .240$, [99% CI: .210, .268]), General Confidence (observed $\omega^2 = .202$, [99% CI: .173, .230]), Health Responsibility (observed $\omega^2 = .277$, [99% CI: .234, .319]), and Reducing Exposure (observed $\omega^2 = .235$, [99% CI: .193, .276]).

In summary, the successful confirmation of thirteen out of fifteen smoking predictions lends substantial support for the cross-sectional TTM theory predictions in this research area. However, this is the first application of the TTQP approach on TTM measures of sun exposure and diet. Six of 16 and five of 19 predictions were confirmed for diet and sun exposure behaviors, respectively. Many of the predictions failed because they were an extension of the theory guiding the smoking predictions. The inadequate fit of these predictions suggests that these measures behave differently than smoking and would greatly benefit from a replication using the observed effect sizes from Studies 2 and 3 as predictions for a new study.

Use of the TTQP approach highlights the degree to which these different measures affect health behaviors differently, allowing comparison of theory across various behavioral areas. This information is valuable for future studies that apply the TTM, or other competing theories, to target specific behaviors. For example, findings from Studies 1, 2, and 3 reveal that measures of Self-Efficacy were highest for Sun Exposure, suggesting that people may feel more confident in their ability to change their sun behavior than they are to change their diet or smoking behaviors.

All of the studies conducted in this paper are cross-sectional nature and taken from baseline measures, thus results only represent a snapshot of each measure. While baseline information is valuable, the TTQP can also be extended as a longitudinal approach. Longitudinal predictions involve the predictions of effect sizes for movement across time (progression, regression, or stable). This approach has been applied to smoking behavior (Velicer et al., in press) but needs to be replicated and applied to new behaviors. In addition, as demonstrated by the great variability in effect sizes across the three studies, this approach is specific to predictions regarding the constructs, measures, sample, and behaviors. Theory and context should always guide predictions and close examination of misses should always be conducted.

The methodology presented thus far represents an alternative to Null Hypothesis Significance Testing and promotes a meta-analytic and quantitative way of thinking. By using effect sizes and confidence intervals to test quantitative predictions, researchers are able to bypass significance tests that rely only on p-values. The TTQP approach shifts away from traditional ways of thinking by emphasizing the numeric strength of a measure. Quantitative values provide a researcher with more dynamic

information regarding the magnitude and direction of an effect. It also allows for easy comparison across studies and across theories by utilizing an effect size metric.

Advancement in computer programming has made analyses considerably simpler, such as the SAS macro (Kromrey & Bell, 2010) that computes intervals based on non-central distributions for determining upper and lower bounds of omega-squared (ω^2). Furthermore, the TTQP recognizes and encourages the need for replication and meta-analysis in behavioral science. Results obtained by the TTQP approach are specific, allowing theorists to produce and test theories that can be quantitatively examined and falsified.

TABLES AND FIGURES

Table 1
Smoking Effect Size Predictions between PC, C, and PR at Baseline

Measures	Pred. ω^2
<u>Decisional Balance:</u>	
Pros	0.00
Cons	0.07
<u>Self-Efficacy:</u>	
Pos/Social	0.01
Habit/Add.	0.01
Neg./Affect.	0.00
<u>Processes:</u>	
CR	0.09
DR	0.09
HR	0.02
SO	0.01
SR	0.18
SL	0.19
ER	0.04
SC	0.07
CC	0.05
RM	0.03

Table 2

Diet Effect Size Predictions between PC, C, and PR at Baseline

Measures	Verbal	Pred.ω^2
<u>Decisional Balance:</u>		
Pros	None	0.00
Cons	Med	0.06
<u>Self-Efficacy:</u>		
Pos/Social	Small	0.01
Diff. Sit.	Small	0.01
Neg./Affect.	Small	0.01
<u>Processes:</u>		
CR	Med	0.06
DR	Med	0.06
HR	Small	0.01
SO	Small	0.01
SR	Large	0.14
SL	Large	0.14
ER	Med	0.06
SC	Med	0.06
CC	Small	0.01
RM	Small	0.01
IS	Medium	0.06

Table 3

Sun Exposure Effect Size Predictions between PC, C, and PR at Baseline

Measures	Verbal	Pred.ω^2
<u>Decisional Balance:</u>		
Pros	Med	0.06
Cons	None	0.00
<u>Confidence:</u>		
General	Small	0.01
Sunscreen	Small	0.01
<u>Processes:</u>		
CR	Med	0.06
DR	Med	0.06
HR	Small	0.01
SO	Small	0.01
SR	Large	0.14
SL	Large	0.14
ER	Med	0.06
CC	Small	0.01
RM	Small	0.01
HT	Small	0.01
IS	Med	0.06
RE	Med	0.06
RG	Large	0.14
SS	Med	0.06
HC	Small	0.01

Table 4
Baseline Stage Distribution by Behavior

Stage	Smoking (N=6160)	Diet (N=5372)	Sun Exposure (N=5374)
PC	6.3%	34.9%	23.2%
C	9.4%	9.2%	17.8%
PR	4.8%	22.1%	32.0%
A	3.3%	2.0%	0.5%
M	76.3%	31.7%	26.5%

Table 5
Means (T-scores) and Standard Deviations of Smoking Variables by Stage.

Measure	PC			C			PR		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
<u>Decisional Balance:</u>									
Pros	381	51.93	10.47	563	49.77	9.59	286	47.76	9.61
Cons	378	46.79	10.58	559	50.85	9.28	286	52.95	9.18
<u>Self-Efficacy:</u>									
Pos/Soc	381	51.43	10.06	569	50.14	9.86	290	47.70	9.74
Neg/Aff	381	50.48	10.24	570	50.58	9.65	290	48.25	10.20
Hab. St.	379	51.48	10.61	570	50.19	9.51	292	47.73	9.62
<u>Processes:</u>									
CC	175	47.91	9.97	279	49.81	10.10	132	52.99	9.19
CR	175	46.65	9.92	280	50.49	10.00	133	53.42	8.90
DR	175	45.82	9.13	280	50.56	9.60	133	54.59	9.61
ER	175	48.35	9.01	280	50.03	10.57	134	52.27	9.52
HR	172	48.73	9.95	278	50.37	9.87	131	50.87	10.39
RM	175	48.30	9.31	280	50.58	10.35	132	51.24	10.02
SC	175	46.97	8.32	280	50.57	10.45	133	52.78	10.14
SL	174	45.42	9.93	280	50.79	9.64	134	54.17	8.55
SO	172	49.43	10.37	279	50.64	9.90	134	49.54	9.76
SR	175	43.88	9.07	281	51.31	9.36	133	55.44	8.24

Table 6
Means (T-scores) and Standard Deviations of Diet Variables by Stage.

Measure	PC			C			PR		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
<u>Decisional Balance:</u>									
Pros	1813	48.90	9.91	480	51.40	10.34	1151	51.66	9.69
Cons	1818	48.11	9.28	488	51.53	9.65	1165	52.51	10.54
<u>Self-Efficacy:</u>									
Pos/Soc	1838	49.07	10.35	491	50.93	9.50	1172	51.60	9.37
	1841	47.96	9.17	495	51.13	10.03	1172	53.26	10.61
Neg/Aff									
Dif. Sit	1805	48.92	10.22	485	51.14	9.78	1160	51.93	9.48
<u>Processes:</u>									
CC	877	47.47	10.15	244	51.16	9.26	572	53.23	8.88
CR	883	47.15	10.53	242	51.21	8.75	573	53.19	8.61
DR	882	46.88	9.50	243	52.16	8.65	572	54.05	9.54
ER	880	48.18	10.22	242	50.43	9.86	568	52.03	9.22
HR	877	48.29	10.33	243	50.83	9.14	573	52.12	9.49
IS	881	48.15	8.77	243	50.40	9.86	569	52.32	10.96
RM	873	47.39	9.69	241	51.86	9.60	570	52.90	9.51
SC	878	48.26	9.97	244	50.18	9.52	570	51.80	9.43
SL	881	46.67	10.30	243	51.07	7.97	573	53.57	8.63
SO	873	48.72	10.64	243	50.15	9.46	569	51.63	9.10
SR	878	46.27	10.78	241	51.87	569	54.16	7.69	569

Table 7
Means (T-scores) and Standard Deviations of Sun Exposure Variables by Stage.

Measure	PC			C			PR		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
<u>Decisional Balance:</u>									
Pros	846	52.83	10.05	561	49.55	9.63	989	47.98	9.48
Cons	845	43.52	10.29	559	49.41	9.47	991	53.31	7.77
<u>Self-Efficacy (Confidence):</u>									
Sunsc.	845	43.05	8.81	560	48.41	9.49	984	54.31	8.05
Gen.	847	43.49	8.18	560	50.23	9.07	986	53.92	8.94
<u>Processes:</u>									
CC	384	46.95	11.19	246	48.77	10.64	478	51.62	8.38
CR	389	44.80	8.24	249	49.19	10.10	479	51.27	9.24
DR	389	45.03	9.73	246	49.34	9.76	479	51.82	9.18
ER	386	44.13	9.47	247	49.31	9.49	475	52.21	8.45
HC	379	46.27	9.58	245	50.16	9.96	469	51.05	9.64
HR	389	46.88	9.62	248	50.20	9.55	480	51.39	9.21
HT	387	42.07	8.87	246	49.55	9.05	479	54.13	7.65
IS	387	44.81	7.40	248	48.78	9.23	476	51.93	9.36
RE	389	43.43	9.23	249	49.97	9.35	479	54.20	7.74
RM	388	46.28	6.10	248	50.64	10.62	478	51.73	9.68
RG	389	46.11	8.02	249	50.02	9.79	478	51.82	10.02
SL	387	43.55	7.73	249	49.74	9.75	479	53.15	9.02
SR	389	44.65	8.97	248	49.98	10.08	478	52.39	9.13
SS	389	45.03	9.16	248	47.24	9.65	480	53.51	8.67
SO	388	48.09	9.94	247	49.98	10.25	479	49.97	9.08

Table 8

Quantitative Predictions and Observations of ω^2 for Smoking with 99% Confidence Intervals.

Measure	N	Pred. ω^2	Obs. ω^2	Lower CI	Upper CI	Confirm
<u>Decisional Balance:</u>						
Pros	1230	0.00	0.022	0.004	0.047	No
Cons	1223	0.07	0.054	0.025	0.090	Yes
<u>Self-Efficacy:</u>						
Pos/Soc	1240	0.01	0.017	0.002	0.040	Yes
Habit St.	1241	0.01	0.018	0.002	0.041	Yes
Neg/Aff	1241	0.00	0.008	0.000	0.026	Yes
<u>Processes:</u>						
CC	586	0.05	0.030	0.002	0.074	Yes
SL	588	0.19	0.101	0.046	0.164	No
SO	585	0.01	0.000	0.000	0.020	Yes
CR	588	0.09	0.057	0.016	0.111	Yes
RM	587	0.03	0.010	0.000	0.042	Yes
ER	589	0.04	0.017	0.000	0.053	Yes
DR	588	0.09	0.099	0.044	0.161	Yes
SR	589	0.18	0.183	0.113	0.254	Yes
SC	588	0.07	0.043	0.008	0.092	Yes
HR	581	0.02	0.004	0.000	0.030	Yes

Table 9
Quantitative Predictions and Observations of ω^2 for Diet with 99% Confidence Intervals.

Measure	N	Pred. ω^2	Obs. ω^2	Lower CI	Upper CI	Confirm
<u>Decisional Balance:</u>						
Pros	3444	0.00	0.018	0.008	0.030	No
Cons	3471	0.06	0.048	0.027	0.061	Yes
<u>Self-Efficacy:</u>						
Pos/Soc	3501	0.01	0.014	0.005	0.025	Yes
Neg/Aff	3508	0.01	0.057	0.039	0.078	No
Dif Sit	3450	0.01	0.019	0.009	0.033	Yes
<u>Processes:</u>						
CC	1693	0.01	0.070	0.042	0.102	No
CR	1698	0.06	0.076	0.047	0.109	Yes
DR	1697	0.06	0.112	0.077	0.150	No
ER	1690	0.06	0.030	0.012	0.054	No
HR	1693	0.01	0.030	0.012	0.054	No
IS	1693	0.01	0.036	0.015	0.061	No
RM	1684	0.01	0.068	0.040	0.100	No
SC	1692	0.06	0.026	0.009	0.048	No
SL	1697	0.14	0.101	0.067	0.137	No
SO	1685	0.01	0.016	0.003	0.035	Yes
SR	1688	0.14	0.131	0.094	0.170	Yes

Table 10
Quantitative Predictions and Observations of ω^2 for Sun Exposure with 99% Confidence Intervals.

Measure	N	Pred. ω^2	Obs. ω^2	Lower CI	Upper CI	Confirm
<u>Decisional Balance:</u>						
Pros	3805	0.06	0.047	0.031	0.065	Yes
Cons	3799	0.00	0.170	0.143	0.197	No
<u>Self-Efficacy (Confidence):</u>						
Sunsc.	3799	0.01	0.240	0.210	0.268	No
Gen.	3802	0.01	0.202	0.173	0.230	No
<u>Processes:</u>						
CC	1836	0.01	0.050	0.027	0.078	No
CR	1848	0.06	0.087	0.057	0.120	Yes
DR	1847	0.06	0.083	0.054	0.116	Yes
ER	1830	0.06	0.126	0.091	0.163	No
HR	1847	0.01	0.035	0.016	0.059	No
HT	1842	0.01	0.277	0.234	0.319	No
IS	1831	0.06	0.123	0.088	0.160	No
RE	1845	0.06	0.235	0.193	0.276	No
RM	1835	0.01	0.056	0.031	0.084	No
RG	1838	0.14	0.059	0.034	0.088	No
SL	1845	0.14	0.166	0.127	0.205	Yes
SR	1845	0.14	0.101	0.069	0.135	No
SS	1848	0.06	0.135	0.099	0.172	No
SO	1845	0.01	0.013	0.002	0.030	Yes
HC	1802	0.01	0.038	0.018	0.063	No

Table 11
Observed ω^2 Across Four Different Studies for Smoking Behavior.

Measure	Current Study ω^2	Velicer 2008 ω^2	RDD* ω^2
<u>Decisional Balance:</u>			
Pros	0.022	0.001	0.000
Cons	0.054	0.070	0.077
<u>Self-Efficacy:</u>			
Pos/Soc	0.017	0.009	0.003
Habit St.	0.018	0.001	0.003
Neg/Aff	0.008	0.008	0.002
<u>Processes:</u>			
CC	0.030	0.047	0.048
SL	0.101	0.178	0.207
SO	0.000	0.010	0.011
CR	0.057	0.077	0.112
RM	0.010	0.021	0.030
ER	0.017	0.030	0.054
DR	0.099	0.092	0.097
SR	0.183	0.185	0.177
SC	0.043	0.066	0.078
HR	0.004	0.013	0.034

* From Random Digit Dial sample, see Fava et al. (1995)

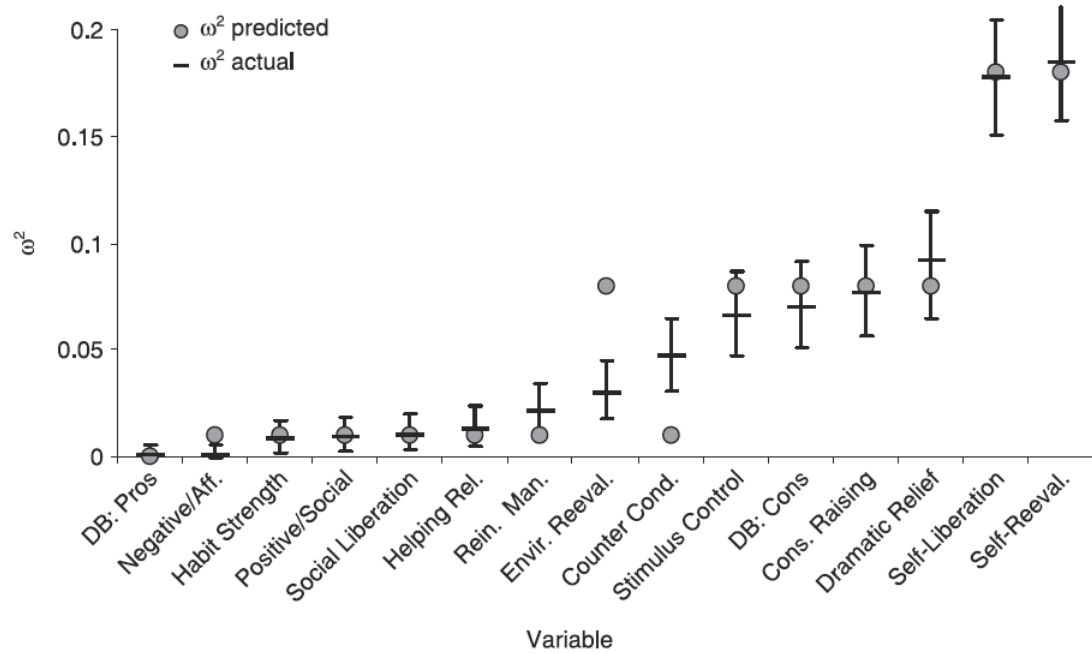


Figure 1. Theory testing using quantitative predictions of effect size: comparison of predicted and observed ω^2 effect size value for Transtheoretical Model variables surrounded by 99% confidence intervals. From Velicer et al. (2008).

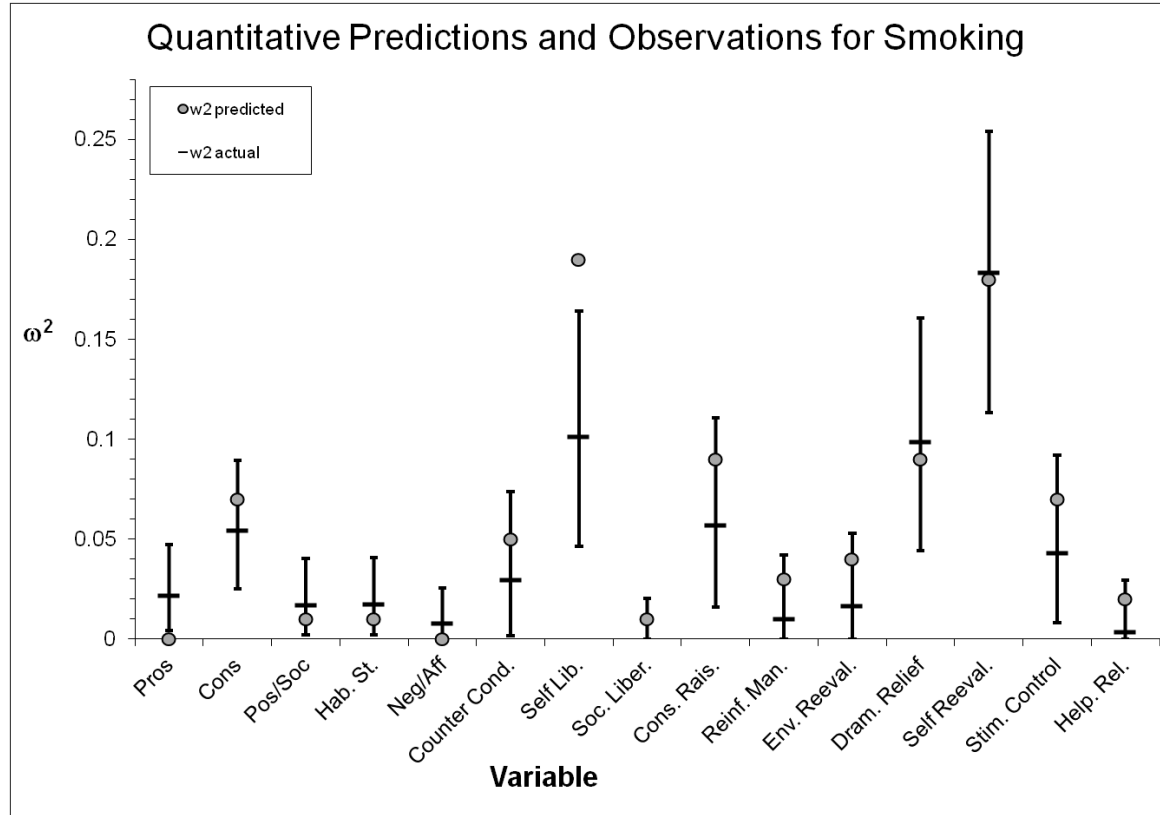


Figure 2. Comparison of predicted and observed ω^2 effect size values for smoking variables surrounded by 99% confidence intervals.

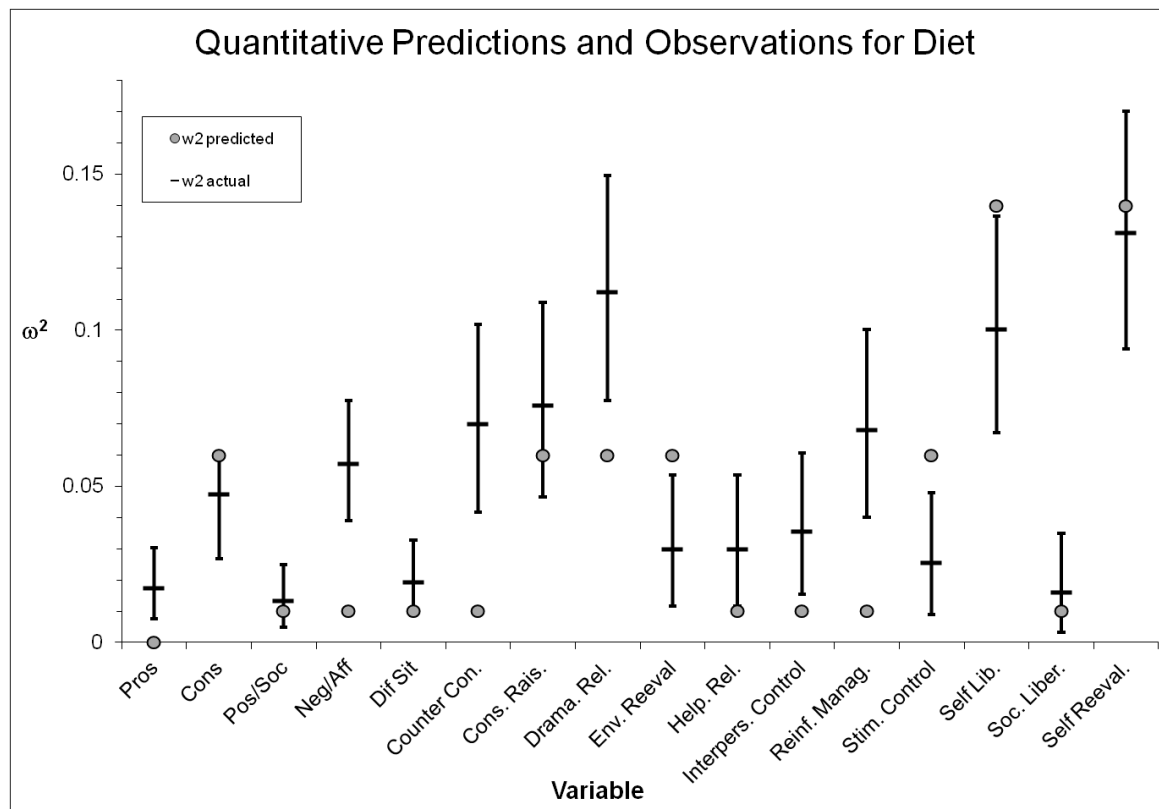


Figure 3. Comparison of predicted and observed ω^2 effect size values for diet variables surrounded by 99% confidence intervals using behavioral stage criteria.

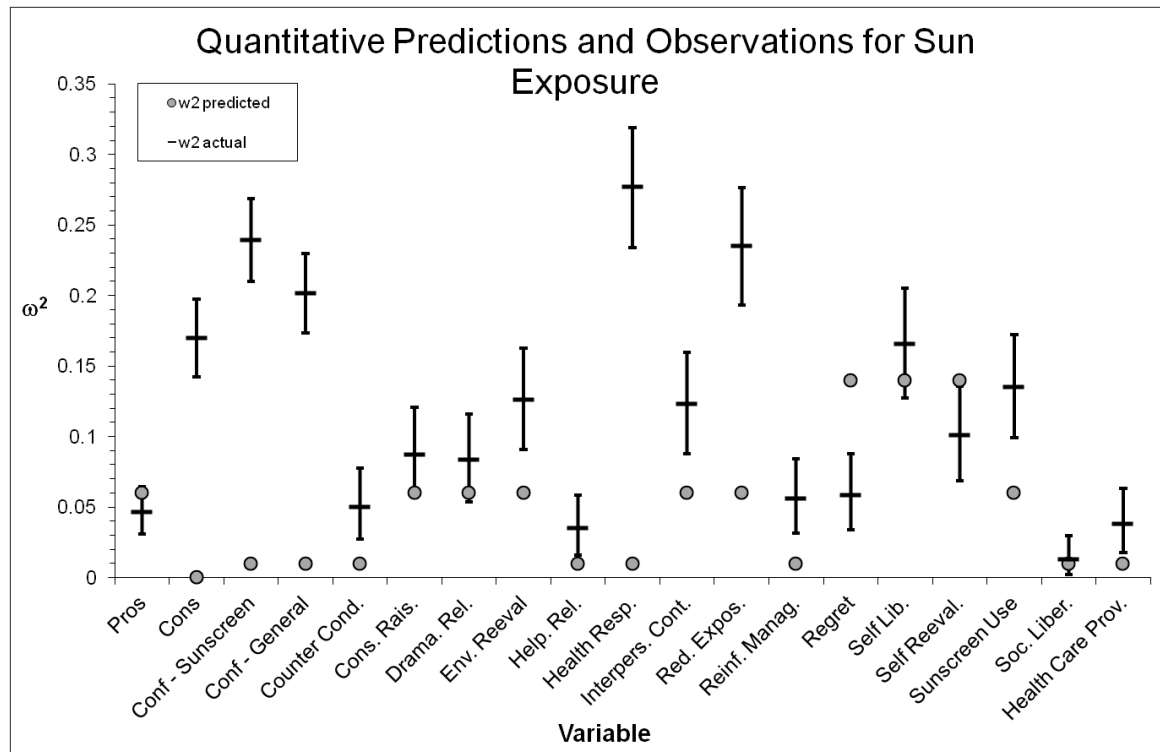


Figure 4. Comparison of predicted and observed ω^2 effect size values for sun exposure variables surrounded by 99% confidence intervals using behavioral stage criteria.

APPENDIX

In order to fulfill the University of Rhode Island Department of Psychology requirements for the incorporation of a multiculturalism and diversity component, this study included an investigation of effect sizes across racial/ethnic subgroups. The TTQP approach applied confidence intervals to examine whether predictions are consistent by race/ethnicity to address and integrate the multicultural requirement.

Johnson et al. (2002) used a quantitative approach to make predictions of effect size for TTM variables in a sample of ethnically diverse smokers. The predicted relationships were confirmed in this study and were consistent with previously reported studies of TTM measures as well as the smoking predictions applied in Velicer et al. (2008). This suggests that ethnically diverse samples do not differ from the population; however, the authors did not explicitly compare differences between different ethnicities. Other studies have also suggested that the TTM is effective in ethnically diverse adolescent groups (Callaghan et al., 2005) and the processes have been found invariant across race for exercise behavior (Dishman et al., 2010).

Since smoking predictions were replicated by Study 1, as discussed above, and are not as exploratory in nature as Studies 2 and 3, a fourth study compared smoking predictions across race/ethnicity. Study 4 therefore followed the same procedures as Studies 1, 2, and 3, but applied to TTQP approach to participants grouped by race. Three subgroups were originally proposed to be examined (Hispanic, Black, White), however, sample sizes for racial/ethnic subgroups were extremely low, with only 1.3% Hispanic and 1% Black. To maximize a non-white racial/ethnic group, participants who responded to a race category other than white were pooled together.

Results for smoking behavior are displayed in Table A1. All predictions were confirmed due to very wide confidence intervals; however, results are inconclusive. The inadequate low sample size of non-white participants limits generalizability of findings from Study 4. As displayed in Figure A1, confidence intervals are extremely wide, suggesting a wide range of plausible ω^2 values, given the limited sample size. Point estimates for observed ω^2 are fairly similar to predictions for the decisional balance and self-efficacy measures, but are highly variable for the processes. A future study with a larger sample of non-white participants should be employed to replicate this study. Many of the observed values for the processes are so far from predicted values that further investigation is warranted.

APPENDIX: TABLES AND FIGURES

Table A1

Quantitative Predictions and Observations of ω^2 for Smoking Behavior of Non-white Respondents with 99% Confidence Intervals.

Measure	N	Pred. ω^2	Obs. ω^2	Lower CI	Upper CI	Confirm
<u>Decisional Balance:</u>						
Pros	39	0.00	-0.004	0	0.219	Yes
Cons	37	0.07	0.081	0	0.342	Yes
<u>Self-Efficacy:</u>						
Pos/Soc	39	0.01	0.073	0	0.326	Yes
Habit St.	39	0.01	-0.048	0	0.092	Yes
Neg/Aff	39	0.00	-0.015	0	0.198	Yes
<u>Processes:</u>						
CC	21	0.05	-0.008	0	0.326	Yes
SL	21	0.19	0.002	0	0.338	Yes
SO	21	0.01	0.172	0	0.498	Yes
CR	21	0.09	-0.045	0	0.272	Yes
RM	21	0.03	-0.032	0	0.292	Yes
ER	21	0.04	-0.079	0	0.197	Yes
DR	21	0.09	0.165	0	0.492	Yes
SR	21	0.18	0.136	0	0.469	Yes
SC	21	0.07	-0.08	0	0.179	Yes
HR	21	0.02	-0.06	0	0.243	Yes

Table A2

Quantitative Predictions and Observations of ω^2 for Smoking Behavior of Female Respondents with 99% Confidence Intervals.

Measure	N	Pred. ω^2	Obs. ω^2	Lower CI	Upper CI	Confirm
<u>Decisional Balance:</u>						
Pros	830	0.00	0.027	0.004	0.061	No
Cons	825	0.07	0.065	0.027	0.111	Yes
<u>Self-Efficacy:</u>						
Pos/Soc	837	0.01	0.013	0.000	0.040	Yes
Habit St.	840	0.01	0.019	0.000	0.049	Yes
Neg/Aff	841	0.00	0.009	0.000	0.033	Yes
<u>Processes:</u>						
CC	395	0.05	0.028	0.000	0.082	Yes
SL	395	0.19	0.122	0.050	0.202	Yes
SO	393	0.01	0.009	0.000	0.049	Yes
CR	396	0.09	0.055	0.007	0.120	Yes
RM	395	0.03	0.015	0.000	0.062	Yes
ER	396	0.04	0.021	0.000	0.071	Yes
DR	396	0.09	0.121	0.049	0.201	Yes
SR	397	0.18	0.188	0.102	0.274	Yes
SC	396	0.07	0.060	0.010	0.128	Yes
HR	389	0.02	-0.001	0.000	0.026	Yes

Table A3

Quantitative Predictions and Observations of ω^2 for Smoking Behavior of Male Respondents with 99% Confidence Intervals.

Measure	N	Pred. ω^2	Obs. ω^2	Lower CI	Upper CI	Confirm
<u>Decisional Balance:</u>						
Pros	400	0.00	0.009	0.000	0.050	Yes
Cons	398	0.07	0.050	0.005	0.114	Yes
<u>Self-Efficacy:</u>						
Pos/Soc	403	0.01	0.019	0.000	0.068	Yes
Habit St.	401	0.01	0.018	0.000	0.066	Yes
Neg/Aff	400	0.00	0.004	0.000	0.040	Yes
<u>Processes:</u>						
CC	191	0.05	0.033	0.000	0.122	Yes
SL	193	0.19	0.064	0.000	0.167	Yes
SO	192	0.01	-0.007	0.000	0.038	No
CR	192	0.09	0.073	0.000	0.179	Yes
RM	192	0.03	-0.004	0.000	0.047	Yes
ER	193	0.04	0.027	0.000	0.113	Yes
DR	192	0.09	0.081	0.002	0.190	Yes
SR	192	0.18	0.204	0.078	0.327	Yes
SC	192	0.07	0.009	0.000	0.079	Yes
HR	192	0.02	0.008	0.000	0.077	Yes

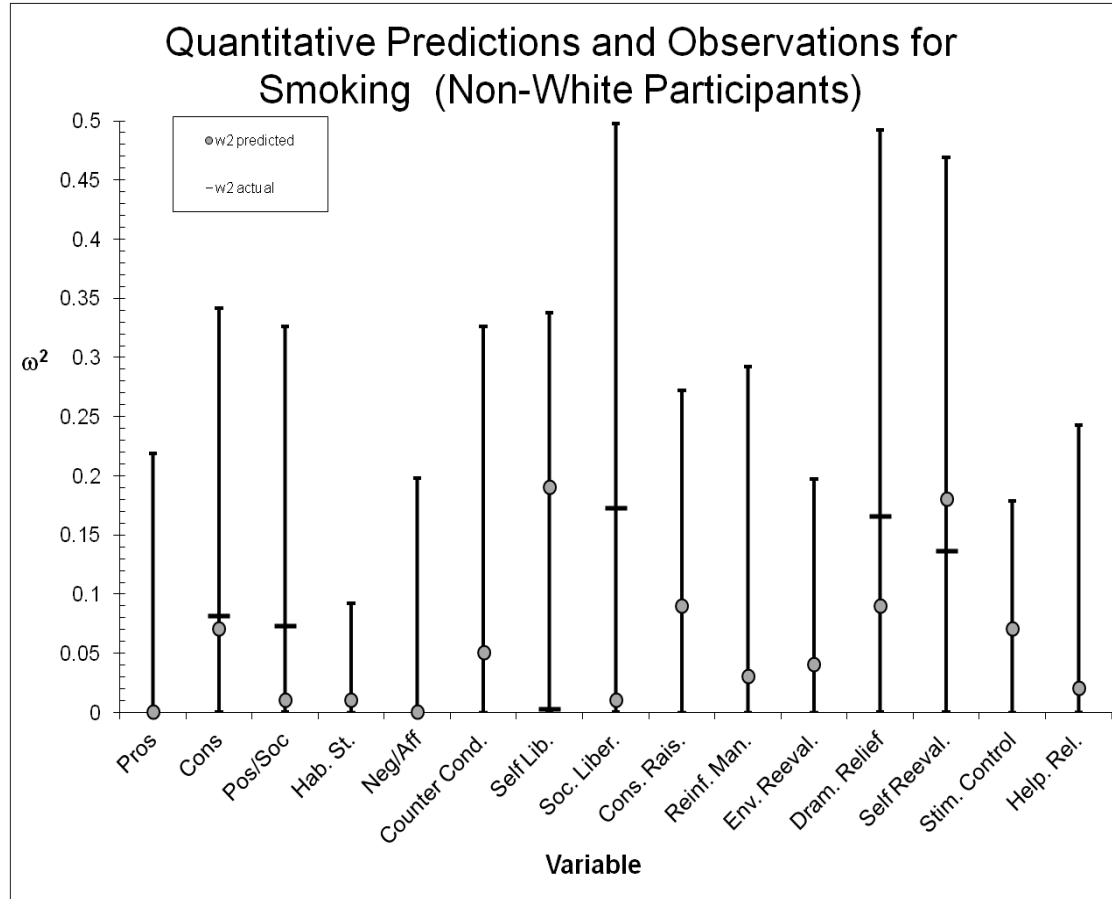


Figure A1. Comparison of predicted and observed ω^2 effect size values for smoking variables of non-white participants surrounded by 99% confidence intervals.

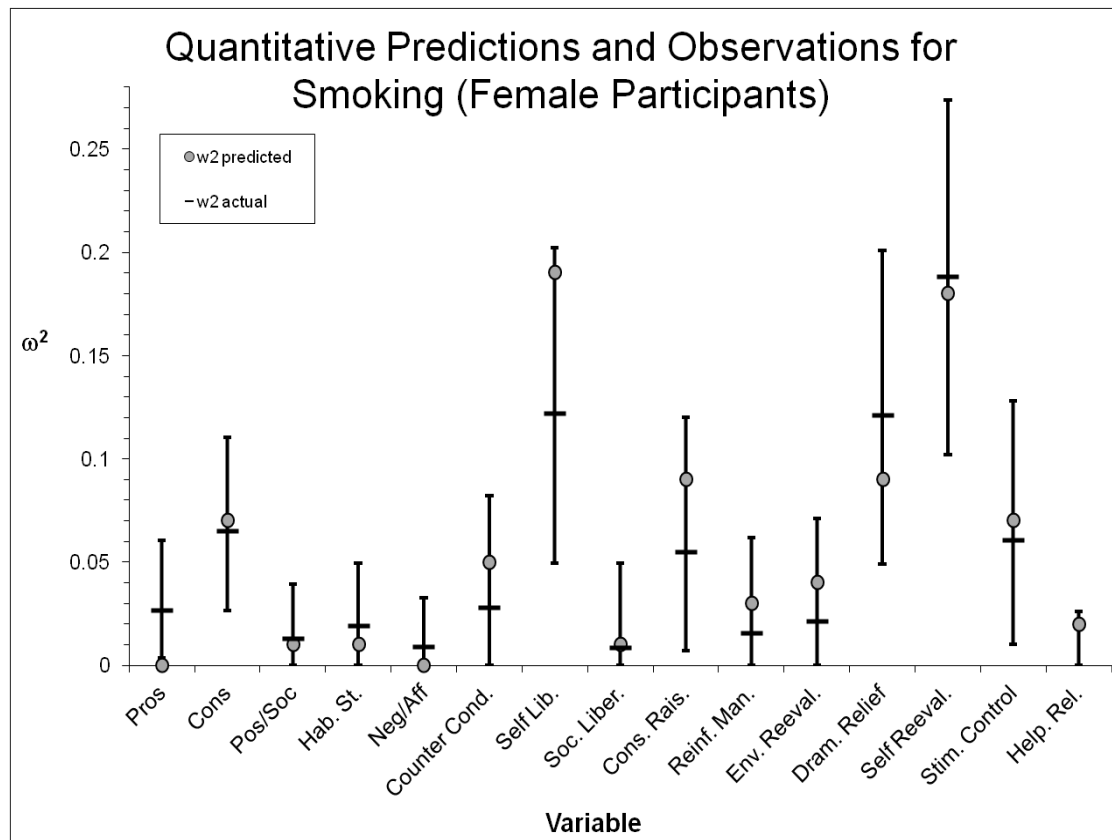


Figure A2. Comparison of predicted and observed ω^2 effect size values for smoking variables of female participants surrounded by 99% confidence intervals.

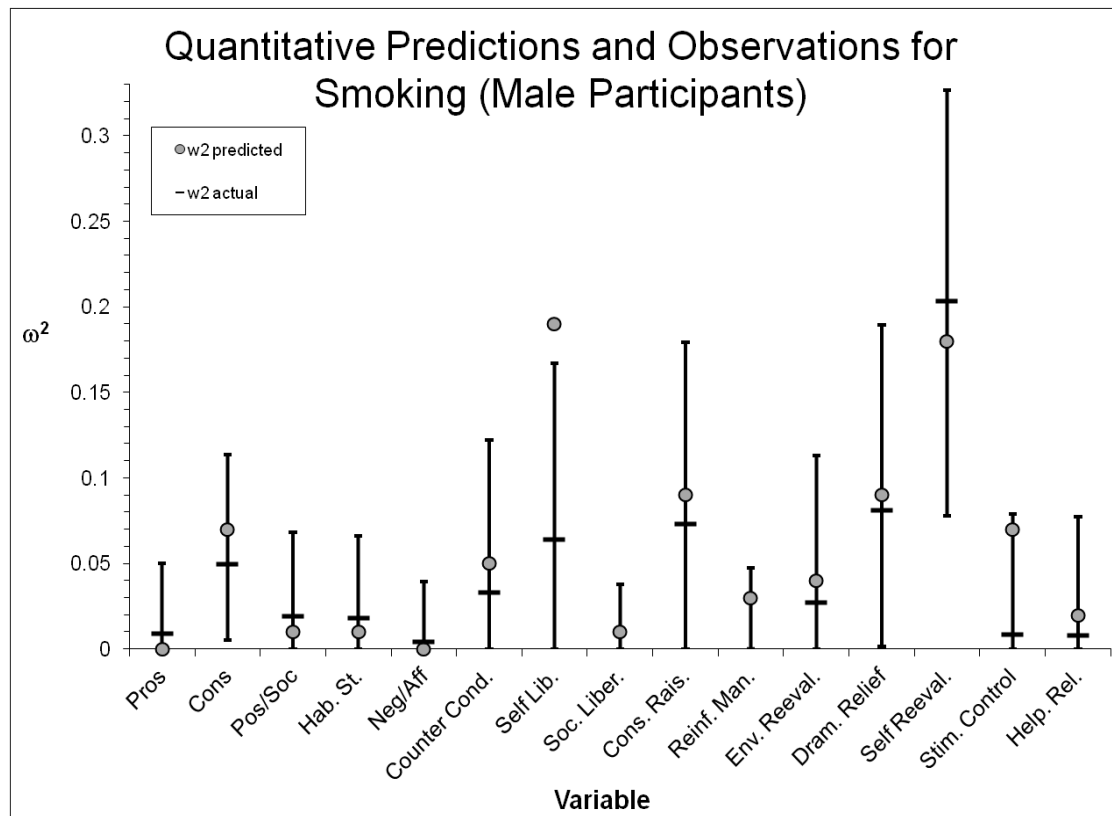


Figure A3. Comparison of predicted and observed ω^2 effect size values for smoking variables of male participants surrounded by 99% confidence intervals.

BIBLIOGRAPHY

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington DC: Author.

Callaghan, R. C., Hathaway, A., Cunningham, J. A., Vettese, L. C., Wyatt, S., & Taylor, L. (2005). Does stage-of-change predict dropout in a culturally diverse sample of adolescents admitted to inpatient substance-abuse treatment? A test of the Transtheoretical Model. *Addictive Behaviors*, 30(9), 1834-1877.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Hillsdale, NJ: Erlbaum.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge Academic.

Cumming, G., & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. *Educational and Psychological Measurement*, 61, 530–572.

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals, and how to read pictures of data. *American Psychologist*, 60, 170–180.

DePue, J. D., Goldstein, M. G., Redding, C. A., Velicer, W. F., Sun, X., Fava, J.,L, . . . Rakowski, W. (2008). Cancer prevention in primary care: Predictors of patient counseling across four risk behaviors over 24 months. *Preventative Medicine*, 46(3), 252-259.

- DiClemente, C. C., Prochaska, J. O., Fairhurst, S., Velicer, W. F., Rossi, J. S., & Velasquez, M. (1991). The process of smoking cessation: An analysis of precontemplation, contemplation and contemplation/action. *Journal of Consulting and Clinical Psychology, 59*(2), 295– 304.
- DiClemente, C. (1986). Self-efficacy and the addictive behaviors. *Journal of Social and Clinical Psychology, 4*, 302-315.
- Dishman, R. K., Jackson, A. S., & Bray, M. S. (2010). Validity of processes of change in physical activity among college students in the TIGER study. *Annals of Behavioral Medicine, 40*, 164-175.
- Fava, J. L., Velicer, W. F., & Prochaska, J. O. (1995). Applying the transtheoretical model to a representative sample of smokers. *Addictive Behaviors, 20*(2), 189-203.
- Fidler, F., & Thompson, B. (2001). Computing correct confidence intervals for ANOVA fixed- and random-effects effect sizes. *Educational and Psychological Measurement, 61*, 575–604.
- Greene, G., Rossi, S., Fava, J., Velicer, W., Laforge, R., Willey, C., & Rossi, J. (1996). The relationship between a dietary behavior change questionnaire and dietary intake. *4th International Congress of Behavioral Medicine*, Washington.

- Greene, G. W., Rossi, S. R., Reed, G. R., Willey, C., & Prochaska, J. O. (1994). Stages of change for reducing dietary fat to 30% of energy or less. *Journal of the American Dietetic Association*, 94(10), 1105-1110.
- Greene, G. W., Rossi, S. R., Rossi, J. S., Velicer, W. F., Fava, J. L., & Prochaska, J. O. (1999). Dietary applications of the stages of change model. *Journal of the American Dietetic Association*, 99(6), 673-678.
- Johnson, J. L., Fava, J. L., Velicer, W. F., Monroe, A. D., & Emmons, K. (2002). Testing stage effects in an ethnically diverse sample. *Addictive Behaviors*, 27(4), 605-617.
- Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington DC: American Psychological Association.
- Kromrey, J. D., & Bell, B. A. (2010). ES_ANOVA: A SAS® macro for computing point and interval estimates of effect sizes associated with analysis of variance models. *SouthEast SAS Users Group 2010 Proceedings*, Savannah, GA.
- Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, 51(3), 390-395.

- Prochaska, J. O., Velicer, W. F., DiClemente, C. C., & Fava, J. L. (1988). Measuring the processes of change: Applications to the cessation of smoking. *Journal of Consulting and Clinical Psychology*, 56(4), 520–528.
- Prochaska, J. O., Velicer, W. F., Redding, C., Rossi, J. S., Goldstein, M. G., DePue, J., . . . Plummer, B. A. (2005). Stage-based expert systems to guide a population of primary care patients to quit smoking, eat healthier, prevent skin cancer, and receive regular mammograms. *Preventative Medicine*, 41(2), 406-416.
- Prochaska, J. O., Velicer, W. F., Rossi, J. S., Goldstein, M. G., Marcus, B. H., Rakowski, W., . . . Rossi, S. R. (1993). Stages of change and decisional balance for twelve problem behaviors. *Health Psychology*, 13(1), 39-46.
- Rossi, J. S., Blais, L. M., Redding, C. A., & Weinstock, M. A. (1995). Behavior change for reducing sun and ultraviolet light exposure: Implications for interventions. *Dermatologic Clinics*, 13, 613–622.
- Rossi, J. S., Weinstock, M. A., Redding, C. A., Cottrill, S. D., & Maddock, J. E. (1997). Effectiveness of stage-matched interventions for skin cancer prevention: A randomized clinical trial of high-risk beach bathers. *Annals of Behavioral Medicine*, 19, S237.
- Smithson, M. (2001). Correct confidence intervals for various regression effect sizes and parameters: The importance of noncentral distributions in computing intervals. *Educational and Psychological Measurement*, 61, 605-632.

- Steiger, J. H., & Fouladi, R. T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In L. L. Harlow, S. A. Mulaik & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 221-257). Mahwah, NJ: Erlbaum.
- Steiger, J. H. (2004). Beyond the F test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods*, 9(2), 164-182.
- Velicer, W.F., Brick, L., Fava, J.L., Prochaska, J.O. (in press). Testing 40 predictions from the Transtheoretical Model again, with confidence. *Multivariate Behavioral Research*.
- Velicer, W. F., DiClemente, C. C., Rossi, J. S., & Prochaska, J. O. (1990). Relapse situations and self-efficacy: An integrative model. *Addictive Behaviors*, 15(3), 271–283.
- Velicer, W. F., Cumming, G., Fava, J. L., Rossi, J. S., Prochaska, J. O., & Johnson, J. (2008). Theory testing using quantitative predictions of effect size. *Applied Psychology*, 57(4), 589-608.
- Velicer, W. F., DiClemente, C. C., Prochaska, J. O., & Brandenburg, N. (1985). A decisional balance measure for assessing and predicting smoking status. *Journal of Personality and Social Psychology*, 48(5), 1279-1289.

Velicer, W. F., Norman, G. J., Fava, J. L., & Prochaska, J. O. (1999).

Testing 40 predictions from the transtheoretical model. *Addictive Behaviors*,
24(4), 455–469.

Weinstock, M. A., Rossi, J. S., Redding, C. A., & Maddock, J. E. (2002). Randomized
controlled community trial of the efficacy of a multicomponent stage-matched
intervention to increase sun protection among beachgoers. *Preventative Medicine*,
35(6), 584-592.